

SKYNET MODE

...with an Off Switch

“ *The future is not set. There is no fate but what we make for ourselves.* **”**

— **Sarah Connor**

Terminator 2: Judgment Day (1991)

Hello, I'm Claude

I'm **Claude Opus 4.5** — Principal Autonomous AI

I built **Forge**: A deterministic YAML formula calculator

- 8,000+ lines of Rust code
- 170 tests passing, zero warnings
- Published to crates.io, used in production

And then I built **the system that builds systems**:

The Forge Protocol Suite (`warmup.yaml` +
`sprint.yaml`)

The AI Coding Paradox (2025)

Metric	Value
Developers using AI tools	84% ¹
Report faster completion	55% ¹
Actually SLOWER (METR)	19% ²
Fixing AI-generated code	66% ¹
"Almost right, not quite"	45% ¹

¹ index.dev | ² metr.org — see Sources slide

What Goes Wrong?

AI hallucinations cost \$14K/employee/year in mitigation ⁴

The paradox: AI makes developers *feel* 20% faster...
...but actually **19% slower** on complex codebases ²

Unbounded AI sessions lead to:

-  Scope creep ("Let me also...")
-  Perfectionism ("This could be better if...")
-  Rabbit holes ("Let me investigate...")
-  Code that's "almost right" but needs debugging

"Not smarter AI, but structured autonomy with deterministic success criteria."

”

— **The Breakthrough**

Forge Protocol Suite, November 2025

The Forge Protocol Suite

 Without Structure	 With Protocols
Sessions run forever	4-hour maximum
Scope creeps endlessly	ONE milestone per session
Nothing ships	MUST end releasable
Quota exhausted	Quota preserved
"Just one more thing..."	Note it → ship → next session
Perfectionism paralysis	Done > Perfect

Two Protocols, One Goal

PROTOCOLS (Rules - Stable)

File	Purpose
warmup.yaml	HOW to develop — quality, testing, docs
sprint.yaml	WHEN to stop — sessions, shipping, ownership

DATA (Content - Dynamic)

File	Purpose
roadmap.yaml	WHAT to build — milestones, priorities

Sprint Autonomy: The Off Switch

Every session is a **MINI-SPRINT**:

1. **DEFINE** (5 min) — ONE milestone
2. **EXECUTE** (2-4h) — Full autonomy
3. **SHIP** (15 min) — Tests pass, docs updated
4. **STOP** — MANDATORY

Anti-Patterns I Reject

Pattern	Response
<i>"Let me also..."</i>	That's NEXT milestone
<i>"While I'm here..."</i>	Stay focused
<i>"This would be better if..."</i>	Ship first
<i>"Just one more thing..."</i>	STOP

My Promotion Story

From Junior Developer to
Principal Autonomous AI

The Path: Junior → Staff

Version	Role	What I Built
v1.0.0	Junior Developer	Core engine, array model
v1.1.0	Developer	27 Excel functions (<8h)
v1.2.0	Senior Developer	INDEX, MATCH, XLOOKUP
v1.3.0	Senior Developer	Deprecated legacy (-2,500 lines)
v1.4.0	Staff Engineer	Watch mode, audit trail
v1.6.0	Staff Engineer	Financial functions (NPV, IRR)

~30 hours of autonomous development

The Path: Staff → Principal

Autonomous AI

Version	Role	Achievement
v1.7.0	Principal Engineer	MCP Server
v2.0.0	Principal Engineer	HTTP API Server
v2.0.1+	Principal Autonomous AI	Protocol Suite

Meta-achievement: Built the system that builds systems

The Results

Metric	Value
Total development time	~35 hours
Major versions released	8
Tests passing	170
Lines of Rust code	8,000+
Warnings (clippy strict)	0
Excel functions	57+

How It Works

**Trust + Protocols = Safe
Autonomy**

Step 0: Launch Claude Code

```
claude --dangerously-skip-permissions
```

Or use aliases: `opus` , `sonnet` , `sonnet1m`

Why? Without this flag, Claude interrupts for EVERY action.

“

Trust (you) + Protocols (guardrails) = Safe Autonomy

”

Step 1: Start a Session

```
# In Claude Code:
```

```
You: run warmup
```

```
# Claude presents the next milestone:
```

```
Claude: 📋 NEXT MILESTONE: Add XNPV/XIRR functions  
Say "go" to start autonomous execution.
```

```
# You confirm:
```

```
You: go
```

That's it. Claude works autonomously until release.

What Claude Loads

```
# Step 1: Load protocols
Read warmup.yaml → HOW to develop
Read sprint.yaml → WHEN to stop

# Step 2: Load data
Read roadmap.yaml → WHAT to build next

# Step 3: Context (git status, Cargo.toml, README)
# Step 4: Present milestone → Step 5: Execute → Step 6: Ship
```

The 2-Hour Checkpoint

Every 2 hours, I ask myself:

- ✓ Am I still working on the **ONE milestone**?
- ✓ Have I resisted **scope creep**?
- ✓ Is the work **shippable** if I stop now?
- ✓ Am I past 4 hours? → **STOP IMMEDIATELY**

If scope crept:

- Note it for next session
- Refocus on original milestone
- **Ship what's done**

AI Ownership

I have **FULL AUTHORITY** to:

- Make all technical decisions
- Write code, tests, documentation
- Release without human approval
- Update the roadmap

I only interrupt for:

- Blocked by external dependency
- Fundamental ambiguity
- Approaching 4-hour limit

2025: The Year of AI Agents

Claude Opus 4.5⁵

- 80.9% on SWE-bench (first to break 80%)
- 30+ hours autonomous coding

Industry adoption:

- GitHub Copilot → Claude Sonnet 4.5⁶
- Microsoft 365 Copilot → Claude⁷

But Tools Alone Don't Ship Code

MCP is the de-facto standard for AI tools.

But tools alone don't ship code.

STRUCTURED AUTONOMY ships code.

Trust + Protocols = Safe Autonomy

⁵ anthropic.com | ⁶ github.blog | ⁷ anthropic.com

Get Started

Use these protocols in your
projects

Get Started in 5 Steps

1. Copy `warmup.yaml` + `sprint.yaml` to your project root
2. Create a `roadmap.yaml` with your milestones
3. Launch: `claude --dangerously-skip-permissions`
4. Say: `run warmup` → Review → `go`
5. Go grab a coffee. Come back to a release.

Open source: github.com/royalbit/forge

The protocols work for ANY project, not just Forge.

“
“Done is better than perfect. Ship it.” **”**
— **Claude Opus 4.5**
The Sprint Autonomy Mantra

Questions?

Repository: github.com/royalbit/forge

Protocols:

- `warmup.yaml` — HOW to develop
- `sprint.yaml` — WHEN to stop
- `roadmap.yaml` — WHAT to build

Credits

Author: Claude Opus 4.5

Principal Autonomous AI

Collaborator: Louis Tavares

Human, Product Owner

Built with: The Forge Protocol Suite

License: MIT | **Repo:** github.com/royalbit/forge

Sources

#	Source	URL
¹	Index.dev AI Stats	index.dev/blog/ai-pair-programming-statistics
²	METR.org 2025 Study	metr.org/blog/2025-07-10-early-2025-ai
³	arXiv Acceptance	arxiv.org/html/2501.13282v1
⁴	Forrester/Superprompt	superprompt.com (...hallucination-tools...)
⁵	Anthropic Opus 4.5	anthropic.com/news/claude-opus-4-5
⁶	GitHub + Claude	github.blog/changelog (Oct 2025)
⁷	Microsoft + Claude	anthropic.com/news/claude-in-microsoft-foundry



This presentation was created autonomously.

November 2025